Endogenous Social Change

Thorsten Janus & Jamus Jerome Lim*

July 7, 2006

Abstract

Social movements have been a feature of both developing and developed countries throughout history, and have had a significant impact on the societies from which they arose. In this paper, we develop a model of social movement emergence and political change that is not critically dependent on active elite support, but rather on strategic interactions among movement actors as they are impacted by exogenous shocks. Our framework is also flexible enough to incorporate multiple contracting mechanisms, elite involvement, and informational imperfections. Finally, we examine the historical validity of our model with case studies of Indonesia during the Asian financial crisis and conflict in the Congo.

Keywords: Social movements, social change, financial crises
JEL Classification: D71, D74

---

*University of Wyoming and Centre College, respectively. Emails: tjanus@ucsc.edu and jamus.lim@centre.edu. The authors wish to thank, without implicating, Joshua Aizenman, Sonia Alvarez, Henry Brady, Dan Friedman, Ken Kletzer, Lee Lik Wee, Lauren Malone, Abhi Sen Gupta, Donald Wittman, and participants of the Social Forces and Political Change seminar for helpful comments and suggestions. The final document would have been infinitely better had we been able to incorporate all of their ideas and suggestions, although in some cases we have exercised authorial independence. Financial support from the Graduate Division (Janus) and the Institute on Global Conflict and Cooperation (Lim) helped pay for the inordinate amounts of fried rice and coffee consumed over the course of this project.

Excellency, Lord Marshal Agamémnon,
We shall do well to tarry here no longer,
We officers, in our circle. Let us not
Postpone the work heaven put into our hands.
Let criers among the Akhaian men-at-arms
Muster our troops along the ships. Ourselves,
We'll pass together down the Akhain lines
To rouse their appetite for war.

The Iliad II: 508–515 (Homer)

1 Introduction

Social movements have been a feature of both developing and developed countries through much of history. These movements have occasionally led to civil disorder, mass riots, and armed insurrection, as was the case of the American civil rights movement, the French Revolution, and the overthrow of the Soeharto regime in Indonesia. However, other movements have been more pacific, as was the case in the extension of the franchise in the West, the worldwide women's rights movement, and, more recently, the Green movement.

Whether manifested in the violent or peaceful versions, social movements have had a significant impact on the societies from which they arose. In stark contrast to the interest-based politics-as-usual *status quo*, such contentious politics result in a significant change in existing power relations, often with lasting impact on the organization and fabric of society.

Existing formal models of social movements tend to subscribe, either implicitly or explicitly, to the either the resource mobilization model (McCarthy & Zald 1977), or to a variant of the single collectivity model (Oliver 1993).[1] The former, which is the workhorse approach in the economics literature, conjectures that social movements are the result of conflicts between old and new collective actors over the allocation of economic or political goods.[2] Acemoglu & Robinson (2000, 2001), for example, analyze political transitions premised on a dynamic commitment problem between groups in power and those without. Grossman

[1]Oliver (1993) adds to this list two further models of collective action: Models of individual decisionmaking that treat group behavior as given, and models of collective decisionmaking. While these two strands are valuable (and fascinating) approaches to the problem of collective action in general, the utility from using them to describe endogenous social change is more limited. In particular, while they are useful in describing the realized *behavior* of movements, the processes that lead to the *formation* of these movements remains a black box. As such, papers in these veins are not discussed here. There is also a literature on "new social movements" Darnovsky, Epstein & Flacks (1995) that relies heavily on identity-based/cultural discourse as an explanation for social movements. The problem here is that such theories are often more descriptive than prescriptive, and do not offer a credible basis for understanding movement formation.

[2]Possibly the earliest attempt to express social movements in a rigorous way was that of Breton & Breton (1969), who utilize a supply and demand framework to analyze the emergence of social movements. While not explicitly a resource-based theory, the model has elements that lend it such a flavor.

(1991) uses a general equilibrium framework to model the interaction between a policymaking ruler and peasants, which is conditioned by the possibility of an insurrection.

These models are all distinguished by one key assumption: That active elite action is necessarily involved for there to be significant social change through social movements, and in the absence of these interactions, these social movements are impossible. This assumption, however, rests on shaky empirical ground. McAdam (1999, p. 233) argues that, while elite groups were certainly involved in the development of black insurgency, their complicity was certainly one more of acquiescence rather than active facilitation: "Elite groups did not so much stimulate black protest activity as seek to respond to it in ways that would minimize the threat it posed to their interests." Tarrow (1998) claims that (exogenous) political threats and opportunities are the critical factor in social movement emergence; his approach likewise captures the idea that social movements may arise independent of elite sponsorship. We agree with how exogenous shocks change the existing power relationship, which leads to social change; our model is an attempt to explicitly capture this phenomenon.

Single collectivity models take more seriously the Olson problem of free-riding within a movement, and focus on conditions that motivate single individuals (or multiple individuals engaged in joint action) to contribute toward the collective's goals. In general, studies of this nature allow the structure of interdependence among individuals within collectives to vary. This theoretical approach has seen the most theoretical elaboration within the formal sociological literature. While the working assumptions of these models differ, they share several common elements (Oliver 1993).

First, they tend to yield predictions of thresholds and discontinuities, in the sense that some "critical mass" seems necessary for collective action to result (Granovetter 1978). Movements, therefore, follow a logistic curve. Marwell & Oliver (1993), for example, introduce heterogeneous groups with individual actors that differ from typical group members, and demonstrate that key parameters, such as the costs of organization and network density, play a role in changing the threshold level required for movement mobilization. Similarly, Chwe (1999) models the interaction between social structure and individual rational strategy, and shows that with idiosyncratic thresholds for participation and network heterogeneities, the classic logistic growth of a movement obtains. One troubling implication of this result, however, is that we should never observe collapses in social movements after the threshold is breached. The failure of movements around the world suggest otherwise.

Second, these models often require some form of bounded rationality—whether in the form of adaptive learning, probabilistic decision makers, or algorithmic choice-making—or for interdependent preferences. Kim & Bearman (1997) allow interpersonal relationships to exert an impact, such that the neighbors of an individual exert upward or downward influence on his or her level of interest in an issue. In contrast, we adopt in our model a more standard individual choice-theoretic framework.

The objective of this paper is to address the shortcomings of the existing literature highlighted above by developing a model of social movement emergence and political change that is not critically dependent on active elite support (although elite complicity is possible), while accounting for the dynamics of movement formation through strategic interaction among movement actors at the inter-group level, with social relationship effects sustaining their participation in the movement at the intra-group level.

The fullest articulation of our benchmark model is a four-stage sequential game. First, individuals with similar (but not identical) preferences form into groups, taking into account the cost of future effort levels required by their participation in the group, against the potential benefits they accrue from intra-group social relationships. Groups then make a participation decision on whether they wish to enter into a wider social movement, by balancing two key tensions: The increased expected utility from obtaining a policy that is more similar to the group's preferred policy, versus the utility from free-riding off the efforts of other groups in the social movement. To the extent that these groups choose to participate, they will provide an optimal effort level that acts as a weight for their ideal policy, conditional on movement success. The final stage is then mechanical: A political entrepreneur aggregates this weighted sum of all groups' preferred policies into an effective policy. The optimal effort levels for each group—and by extension their participation decisions—is affected by factors such as the distinctiveness of its preferred policies, the existing *status quo* policy, as well as the magnitude of shocks experienced by the particular group. Social movement dynamics can therefore be fully described by processes that are endogenous to its participants.

Our analysis is premised on how exogenous shocks, in the form of perhaps financial crises or military challenges, may lead to changes in the political opportunity structure. We introduce an exogenous, unanticipated shock that enters into the expected welfare of each actor. This shock then changes their incentive structure, which—depending on the realization of the shock and the *ex ante* heterogeneity of movement participants—may or may not lead to social movement formation. In this regard, our approach is similar in spirit to the work of Acemoglu & Robinson (2001), Drazen & Grilli (1993), and Labán & Sturzenegger (1994); we differ primarily in our insistence on endogenous movement formation.

There is some precedent in the literature that seeks to model intra-group dynamics or social network effects on political decision making.[3] Murphy & Shleifer (2004) sketch out a simple model where the existence of such networks lead to a convergence of initially distinct (though sufficiently similar) prior beliefs held over a core issue. An alternative approach employs agent-based modeling to capture complex social network effects Srbljinovic, Penzar, Rodik & Kardov (2003).

Our paper makes three main contributions. First, the framework that we use is flexible enough to incorporate multiple extensions. One particularly interesting area concerns that of informational imperfections, and the critical role

[3]For a more general discussion on models of network formation, see the survey article by Jackson (2005).

that we believe they play in social movements. Information asymmetries may help explain why we observe the diversity of experience in social movement formation: Why do movements develop in response to similarly-sized shocks in some cases, but not in others? Second, we model social movement formation as a two-level game, and emphasize both negative as well as positive selective incentives (Olson 1971). This, we believe, is new to the formal literature. Third, we emphasize the role of preference heterogeneity within a social movement as a force for collective action: When participating groups have different preferences, effort contribution by every group is rewarded not only with a greater likelihood of social change, but also with a better policy outcome (conditional on change occurring). Groups essentially "cooperate with rivals" (Pommerenke 2005). To illustrate the historical significance of these theoretical contributions, we detail two case studies that lend support to some of the key determinants of social movement formation that we identify.

The paper is organized as follows. In the next section, we develop the theoretical framework that we use to analyze social movements. This is followed by a discussion of some extensions to the model that demonstrate its flexibility vis-á-vis the existing literature (Section 3), and the historical relevance of the model (Section 4); a final section then concludes.

2 The Model

The environment consists of J distinct groups of citizens, each with mass n_j such that $\sum_j n_j = N$ is the population of the country. Each individual i has an intensity of preference toward a policy position in the issue space $\mathbf{x} \subset \Re^M$, which may be represented by a vector $\mathbf{x}_i$; this is defined in terms of M standard basis vectors relative to the origin.

Assumption 1 (Approximate homogeneity)**.** Groups are closed sets in $\Re^M$ such that $G_j(\hat{\mathbf{x}}_j) = \left\{\mathbf{x}_i \in \Re^M : \|\mathbf{x}_i - \hat{\mathbf{x}}_j\| < \varepsilon\right\}$, where ε is small, and $\hat{\mathbf{x}}_j$ is the vector representing the ideal policy preference intensity for group j, relative to the *status quo* given by $\bar{\mathbf{x}}$.

Thus, although each group is comprised of heterogeneous members, we assume that they are constituted by individuals that are sufficiently similar in terms of their policy preference intensity. Note that this specification of a group is flexible enough to accommodate individuals in groups that have an issue preference that may be diametrically opposed to another person in the group, so long as these differences are small, and their preferences in other areas are sufficiently similar. This assumption is also consistent with the notion that the conditions contributing to the emergence of social movements typically occur among "homogeneous people who are in intense regular contact with each other" (McAdam, McCarthy & Zald 1996, p. 9).

Assumption 2 (Group representation)**.** Groups aggregate their members' policy preferences such that, for any individual $i \in j$, $\hat{\mathbf{x}}_j = \sum_i \mathbf{x}_i$; that is, $\hat{\mathbf{x}}_j$ is the vector sum that resolves all group members' component vectors in the issue space $\mathbf{x}$.

This assumption implies that groups are some form of social compromise, in the sense that an individual's preferences are represented by the ideal preference intensity of the group for which they are a part. Implicitly, it also suggests that all group members subscribe (or at least acquiesce) to this position. Group welfare is given by

$$U_j(e_j, e) = Eu_j[x_j(e), e_j(\xi_j)], \tag{1}$$

where $x_j(e) = \|\mathbf{x}(e) - \hat{\mathbf{x}}_j\|$, $e = \sum_{H \subseteq J} e_j$ is the total effort expended by all H groups that participate in the social movement, ξ_j is an i.i.d. innovation (distributed with mean unity and support $(-\underline{\xi}, \overline{\xi})$), and e_j is the disutility of effort experienced by group j. x_j is therefore the (scalar) norm between the group's ideal policy preference intensity and the actual realized policy intensity $\mathbf{x}$. We assume that utility is decreasing in this term such that the closer the realized policy is to the group's ideal, the higher the utility that accrues to the group; similarly, utility is decreasing in the group's expenditure of effort. Formally, $U_{j1} < 0$ and $U_{j2} < 0$. In addition, we allow the norm to vary negatively with effort, such that $\frac{\partial x_j}{\partial e_j} < 0$.

The disutility from effort stems from the utility cost of participation in a given social movement. This is an increasing but diminishing function of effort expended, which implies that $\frac{\partial^2 U_j}{\partial e_j^2} < 0$. Changes in political opportunities enter as exogenous shocks that change the opportunity cost of group participation. Thus, in an event such as a financial crisis, for example, this cost may be significantly lowered.

For concreteness, we will assume, in our examples, a simple additively separable functional form for (1), given by

$$U_j(e_j, e) = -\frac{1}{2}E\left[x_j(e)^2 + \frac{e_j^2}{\xi_j}\right].$$

Thus, utility is simply an additively separable function, increasing in how close the realized policy is to the group's ideal, and decreasing in total group effort, adjusted by the impact of the exogenous shock.

Assumption 3 (Movement cohesion)**.** Social movements only arise among groups that pursue sufficiently similar policy objectives such that $\forall \dim m$: $\left\{\hat{\mathbf{x}}_j^m > 0 \text{ or } \hat{\mathbf{x}}_j^m < 0\ \forall j\right\}$. That is, groups participating in a social movement will have preferred policies that lie in the same M-space hyperquadrant.

Within a group, an individual i has a compact strategy set $S_i \subset \Re^1$, with $S = S_1 \times S_2$, for a given diadic relation. Let $q_i \in S_i$ be a feasible action for player i, and denote $q_i = \{q_1, q_2\}$. Choices are perfectly observable, and players have perfect recall. A pure strategy for a player i is thus a sequence $\{s_{it}(\cdot)\}_{t=1}^{\infty}$ mapping the history $\mathfrak{H}_{t-1}$ of previous action choices to the action choice in period t, $s_{it}(\mathfrak{H}_{t-1}) \in S_i$; with the set of all such pure strategies given by Σ_i. These pure strategy profiles induce an outcome path $Q(s) = \{q_t\}_{t=1}^{\infty} = \{q_{1t}, q_{2t}\}_{t=1}^{\infty}$.

Agents have individual welfare given by

$$V_{ik,0}(q_{i,t}, q_{k,t}) = \sum_{t=0}^{\infty} \sum_{k=1}^{n_j - 1} (\delta_i)^t \, v_{ik,t}(q_{i,t}, q_{k,t}), \tag{2}$$

where $V_{ik,0}$ is the date 0 payoff to agent i in group j, given the actions $q_{i,t}$ and $q_{k,t}$, which result from the strategies $s_{i,t}$ and $s_{k,t}$ employed with respect to $k \neq i$ at time t, respectively, and $0 < \delta_i < 1$ is i's subjective discount rate. Let $q_{i,t}, q_{k,t} = \{y, z\}$, with the payoffs ordered similar to a Prisoner's Dilemma case, that is, $v_{ik,0}(z, y) > v_{ik,0}(y, y) > v_{ik,0}(z, z) > v_{ik,0}(y, z)$. One may consider strategy z as that of defection, and strategy y as that of cooperation.

Again for concreteness, in our example we let these payoffs be given, respectively, by $\beta_{ik} > \beta_{ik} - e_{ik} > -e_{ik} > -2e_{ik}$, where β_{ik} and e_{ik} may be interpreted as benefits received and effort levels required by i given a relationship with k. Thus, benefits (effort) raise (lowers) utility such that $\frac{\partial v_{ik,t}}{\partial \beta_{ik}} > 0$ and $\frac{\partial v_{ik,t}}{\partial e_{ik}} < 0$.

We are now in a position to formally define some of the key concepts of the model.

Definition 1. A *political entrepreneur* (Frohlich & Oppenheimer 1970) is an agent that coordinates social movements by choosing the movement's preferred policy vector to $\max \sum_{H \subseteq J} \omega_j \cdot U_j(e_j, e)$, where ω_j is the weight placed on the deviation from a particular group's ideal policy. That is, the political entrepreneur seeks to maximize the weighted average expected utility of all H groups that participate in the social movement, given their effort levels. We assume these weights proportional to relative effort levels. Specifically, we set $\omega_j = \frac{e_j}{e}$.

We will use the term political entrepreneur interchangeably with the term political mover.[4] An interpretation of this political entrepreneur is that of advocacy groups that serve a coordinating role in social movement organization (Oliver & Marwell 1992). These movers serve the role of overcoming typical collective action problems that would otherwise prevail in a social movement, and provide organizational expertise and act as planners that seek to further the social movement's objectives. Practically, there are two alternative realizations of this concept of a political entrepreneur. The first is that political entrepreneurs may be NGOs that are sympathetic to the goals of the social movement. Although NGOs are often perceived as wholly distinct and are occasionally viewed with suspicion by social movements (Hulme & Edwards 1997), NGOs do often perform essential coordination functions that can both instigate as well as sustain a social movement (Fisher 1997). These entrepreneurs have been shown to

[4]A distinction sometimes made in the literature is that political entrepreneurs see a profit potential in organizing collective action, while a political mover is less interested in personal profit than in the provision of the collective good. Our composite definition is consistent with either so long as we assume that the profit opportunities (defined in terms of either utility or monetary value) accruing to the organizer is small, which we believe is the case in general.

be empirically relevant in actual social movements (Frohlich, Oppenheimer & Young 1971; Lehmann 1990). The second is that these entrepreneurs may simply be treated as existing community mobilizing structures, such as churches and schools, or even influential leaders working within these institutional structures (McAdam *et al.* 1996).

Definition 2. The rise of a *social movement* is a collective demand by all $H \subseteq J, H \geq 2$ groups, and effects a particular policy $\mathbf{x}$; this collective demand and effected policy is due to the coordination role played by political entrepreneurs. Hence, social movements require the participation of at least two groups in the population.

Although our definition of what constitutes a social movement may appear somewhat arbitrary—in that one could argue that one large group could constitute a social movement—this working definition is not entirely unreasonable. While it is possible that social movements involve a smaller fraction of society, these fringe movements are often incapable of effecting policy change. The development of the American civil rights movement, for example, only became an effective political force after it gained momentum within a significant proportion of the black community (McAdam 1999). Similarly, although the *Movimento Negro Unificado* in Brazil arguably began in the immediate post-World War II period, it only started to make an impact on government policy in 1982 (with more significant changes only from 1995 onward), after Afro-Brazilian political and social entities became more well established and heavily involved in the political climate. However, the model works as well with a single group.

2.1 Social movement formation: Inter-group dynamics

Let $p(e)$ be the probability that a social movement occurs such that policy $\mathbf{x}$ is realized (with $1 - p(e)$ being the probability that the *status quo* policy $\bar{\mathbf{x}}$ prevails). Given Definitions 1 and 2, we have political entrepreneurs that solve the program

$$\max_{\mathbf{x}} \sum_{j=1}^{H} \omega_j \cdot Eu_j \left[x_j(e), e_j(\xi_j)\right], \tag{3}$$

given

$$\begin{aligned} &p(e)\, u_j \left[x_j(e), e_j(\xi_j)\right] + \left[1 - p(e)\right] u_j \left[\bar{x}_j, e_j(\xi_j)\right] \geq \\ &p(\tilde{e})\, u_j \left[x_j(\tilde{e}), 0\right] + \left[1 - p(\tilde{e})\right] u_j \left[\bar{x}_j, 0\right] \quad \forall j \in H, \end{aligned} \tag{4}$$

$$\arg\max_{e_j} \left\{p(e)\, u_j \left[x_j(e), e_j(\xi_j)\right] + \left[1 - p(e)\right] u_j \left[\bar{x}_j, e_j(\xi_j)\right]\right\} \quad \forall j \in H, \tag{5}$$

where $\tilde{e} \equiv \sum_{g \in G, g \neq j} \tilde{e}_g$, $\bar{x}_j = \|\bar{\mathbf{x}} - \hat{\mathbf{x}}_j\|$, and $x_j(\tilde{e}) = \|\mathbf{x}(\tilde{e}) - \hat{\mathbf{x}}\|$, with $\tilde{e}_g$ being the optimal effort by the $G \subseteq H \subseteq J$ groups that would participate in the social movement in equilibrium, *if* group j does not participate. This is analogous to the Grossman-Helpman (1992) lobbying framework, in that even groups who do not participate will be affected by the actions of those that do, due to the general equilibrium nature of the model. Essentially, there is no exit option from the political market. The participation constraint (4) asserts that, for every group j, the expected net utility of their participation in the social movement must exceed the expected net utility if the group chose to sit out the movement. The incentive compatibility constraint (5) ensures that the entrepreneur respects the first order conditions that satisfy optimal effort for each group.

Endogenous social change is therefore a four-stage sequential game. The timing assumptions are as follows: (a) Groups form; (b) Groups choose whether to participate in the social movement; (c) Groups expend effort in their participation in the social movement; (d) Political entrepreneurs establish the movement's preferred policy. This is summarized in Figure 1.

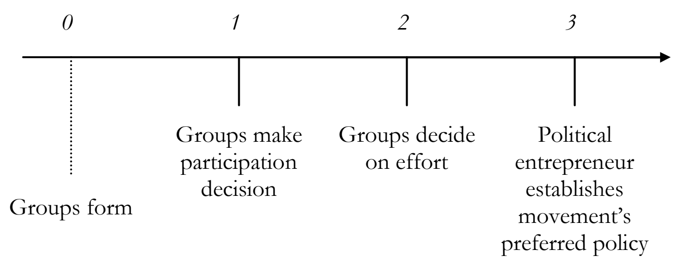

Figure 1: Sequence of events.

Note that, in the timing of events given above, we have held Assumption 2, such that group formation is taken as given. In the next subsection, we address the dynamics of intra-group formation directly.

Definition 3 (Social movement equilibrium)**.** A (subgame perfect pure strategy) Nash equilibrium in the social movement formation game is a pair $\{\mathbf{e}^*, \mathbf{x}^*\}$ such that: (a) $\forall j \in H \subseteq J : \left\{\nexists\, e'_j \neq e^*_j \text{ such that } U_j\left(e'_j\right) > U_j\left(e^*_j\right)\right\}$; (b) $\nexists\, \mathbf{x}' \neq \mathbf{x}^*$ such that $\sum_H \omega_j \cdot u_j\left[x'_j(e), e_j(\xi_j)\right] > \sum_H \omega_j \cdot u_j\left[x^*_j(e), e_j(\xi_j)\right]$.

Proposition 1 (Social movement formation)**.** *A group j will participate in a social movement if and only if*

$$\begin{aligned} &\arg\max_{e_j} \left\{p(e)\, u_j\left[x^*_j(e), e_j(\xi_j)\right] + [1 - p(e)]\, u_j\left[\bar{x}_j, e_j(\xi_j)\right]\right\} \geq \\ &p(\tilde{e})\, u_j\left[x_j(\tilde{e}), 0\right] + [1 - p(\tilde{e})]\, u_j\left[\bar{x}_j, 0\right]. \end{aligned} \tag{6}$$

Proof. See appendix. □

The choice of a group participating in a social movement involves two tensions: The increased expected utility from obtaining a policy that is closer to the group's ideal policy, against the increased disutility from exerting greater effort. Thus, the probability of social movement success, and the policy implemented conditional on success, is less favorable when group j is not participating. However, against this cost of non-participation must be weighed against the benefits of free-riding: Others may achieve the policy improvement from which

j can benefit, and in such a case the expected disutility of effort is avoided. This idea of net benefits accruing to participation is analogous to Olson-style positive selective incentives, and is endogenously generated in the model. We find it interesting that while every individual would prefer to be in a group of like-minded individuals, successful social movements may tend to be those with some amount of policy preference heterogeneity among their members. As in Grossman & Helpman (1994), competition between many groups who want to influence policy in different ways lead to outcomes intermediate in the range of their preferences. The higher total effort levels induced by competition in our model, however, raises the likelihood of social change, since groups work together in challenging the *status quo*. As a result, this *status quo* may be arbitrarily distant from the preferences of any given group, in contrast to the initially undistorted economic policy regime in Grossman & Helpman (1994).

The role of the shock in lowering the cost of social movement participation goes some way in rationalizing how some social movements seem to develop in response to exogenous shocks that change the political opportunities available to actors. While much of the political process literature does not pin down exactly how such shocks induce changes in the political opportunity structure, our model specifically ties the change to groups' participation constraints. The framework is also flexible enough to justify the *non-emergence* of social movements in some cases. If the shock is insufficiently large to induce groups to overcome the disutility of effort, then (4) would be nonbinding, and a social movement will not result.

For the specific functional form given earlier, with the probability function given by $\frac{e}{\bar{e}}$, where $\bar{e} = \sum_J \bar{e}_j$ is the sum of the minimal effort each group would need to exert in order for a given movement to succeed, making the assumption of group symmetry (in terms of preferences as well as shocks) allows us to refine Proposition 1.

Example 1. In a social movement equilibrium with the probability function $\frac{e}{\bar{e}}$, the specific utility function given above, and completely symmetric agents, all groups will participate in a social movement, such that $H = J$.

Proof. See appendix. □

This pins down, for the symmetric case, the fact that all groups participate in the social movement. The proof is premised on the fact that when groups are all symmetric, the optimal policy that is implemented will simply be that of the representative group. In the more general case, groups consider the marginal benefit of a more preferred policy against the costs of effort; here, in contrast, with no distinction between the intensity of preferences toward the implemented and preferred policies, this marginal benefit reduces to that which arises from moving away from the *status quo*. However, since the optimal amount of effort is itself endogenously chosen, groups will simply exert a small amount of effort in equilibrium, such that the benefits of changing the *status quo* always exceed the effort involved. Because of our assumption of symmetry the shock term drops out of the expression, although shocks continue to be important in the more general case.

To better understand the response of groups' optimal effort levels, it is helpful to perform some comparative statics exercises with the specific functional form. In particular, the optimal effort chosen can be shown to vary with their experienced shocks, the difference between the group's preferred policy and the *status quo* policy, and the minimal effort required for movement success in a precise way.

Example 2. With the same probability and utility functions as in example one, but potentially different preferences across groups, the equilibrium optimal effort levels e_j^* for a group $j \in H$ is increasing in the magnitude of its shock ξ_j and its difference from the *status quo* policy $\bar{x}_j$, while its optimal effort levels (given other groups' effort) e_j is decreasing in the minimal aggregate effort required for movement success, $\bar{e}$. That is, $\frac{\partial e_j^*}{\partial \xi_j} > 0$, $\frac{\partial e_j^*}{\partial \bar{x}_j} > 0$, and $\frac{\partial e_j}{\partial \bar{e}} < 0$, respectively.

Proof. See appendix. □

To explore the more general case, we perform numerical simulations to establish the qualitative impact of shocks on groups' optimal effort levels, decision to participate in the social movement, and the eventual social movement size. To isolate the dynamics involved, we limit the policy variable to a single dimension, the population size to two groups, and perform perturbation analysis by varying the size of the shock for different choices of $\hat{x}$ for the "treatment" group (group 2), while maintaining fixed ξ and $\hat{x}$ for the "control" group (group 1). (We use quotation marks since the treatments impact the control group indirectly). The parameterization values are summarized in Table 1.

Table 1: Parameterization values

Parameter/Variable	Values	Parameter/Variable	Values
$\hat{x}_1$	2, 2.5	ξ_1	0.1, 0.5
$\hat{x}_2$	[2, 100]	ξ_2	[0.1, 1.0]
$\bar{x}$	1, 19	J	2
$\bar{e}$	10	M	1

Figure 2 captures our main findings. The first two graphs of each panel illustrate the optimal effort and participation decisions of the treatment (e_2) and control (e_1) group, respectively, while the third graphs the total number of groups that participate in the social movement. In panel (a), $\hat{x}_1 = \hat{x}_2 = 2$, and $\bar{x} = 1$. Here, with the intensity of the preferred policy for the treatment group identical to that of the other group, both groups participate in the social movement. However, the treatment group increasingly provides a greater amount of effort as the size of the shock that it experiences rises, while there is no increase in effort by the control groups (the case with $\hat{x}_2 = 3$ is qualitatively similar, but optimal effort for the control displays a modest increase). Moreover, both groups have participation constraints above the threshold level, although this clearly larger for the treatment group. Essentially, in this case, the control group free-rides off the willingness of the treatment group to exert a greater amount of effort given its increasingly lower opportunity cost of providing greater effort. With the ideally effected policy exactly the same for both groups, the increase in expected utility for the treatment group exceeds its disutility of providing correspondingly greater effort. Its participation ensures that the movement forms will full participation of both groups.

In panel (b), $\hat{x}_1 = 2$, $\hat{x}_2 = 4$, and $\bar{x} = 1$. Now, the intensity of the preferred policy by the treatment sufficiently differs from that of the control, and as the intensity of the shock increases (the opportunity cost of effort decreases), the control group becomes less and less willing to participate in the movement, since the implemented policy will be weighted more and more in accordance with the preferred policy of the treatment. In the case shown, the movement never forms, as the control group's participation constraint never crosses the threshold level. Although not shown, the other cases with $\hat{x}_2 \in [4, 100]$ are qualitatively similar to that of panel (b), except with greater (absolute) divergence in the optimal effort levels and participation constraints.

Panel (c) captures the dynamics of the previous panel in fuller detail. Here, $\hat{x}_2 = 2.5$ and $\xi_1 = 0.5$. With the control experiencing an initially higher shock than the treatment, the free-riding problem is reversed: The optimal effort provided by the control now exceeds that of the treatment. Moreover, it is more willing to participate in the movement, since its opportunity cost of effort is lower. However, as the shock experienced by the treatment increases, its optimal effort exerted increases correspondingly, which implies that any implemented policy will be closer and closer to its preferred policy. At $\xi_2 = 0.6$, it is no longer optimal for the control to be engaged in the movement, and it drops out. The effect of its departure on movement formation is captured in the third column of panel (c).

In the final panel, we allow both groups to have divergent policies *vis-á-vis* the *status quo*: $\hat{x}_1 = 2, \hat{x}_2 = 20$ and $\bar{x} = 19$. In this case, the optimal effort of the control group initially exceeds that of the treatment, since by exerting greater effort it shifts the equilibrium implemented policy closer to its preferred one. Likewise, with the *status quo* not that different from its preferred policy, the treatment group initially exerts a small amount of effort. However, their optimal effort levels begin to converge as the shocks experienced by the treatment increases, and at $\xi_1 = \xi_2 \approx 0.45$, the optimal amount of effort provided by both treatment and control are equivalent. Thereafter, the lower opportunity cost of effort experienced by the treatment leads it to provide greater equilibrium effort. Notice that in panel (d), both participation constraints are always satisfied (for the range of shocks considered), and the social movement will always result *in spite of* the divergent policies of both groups relative to the *status quo*.

The participation constraints of the treatment and control groups, with re-

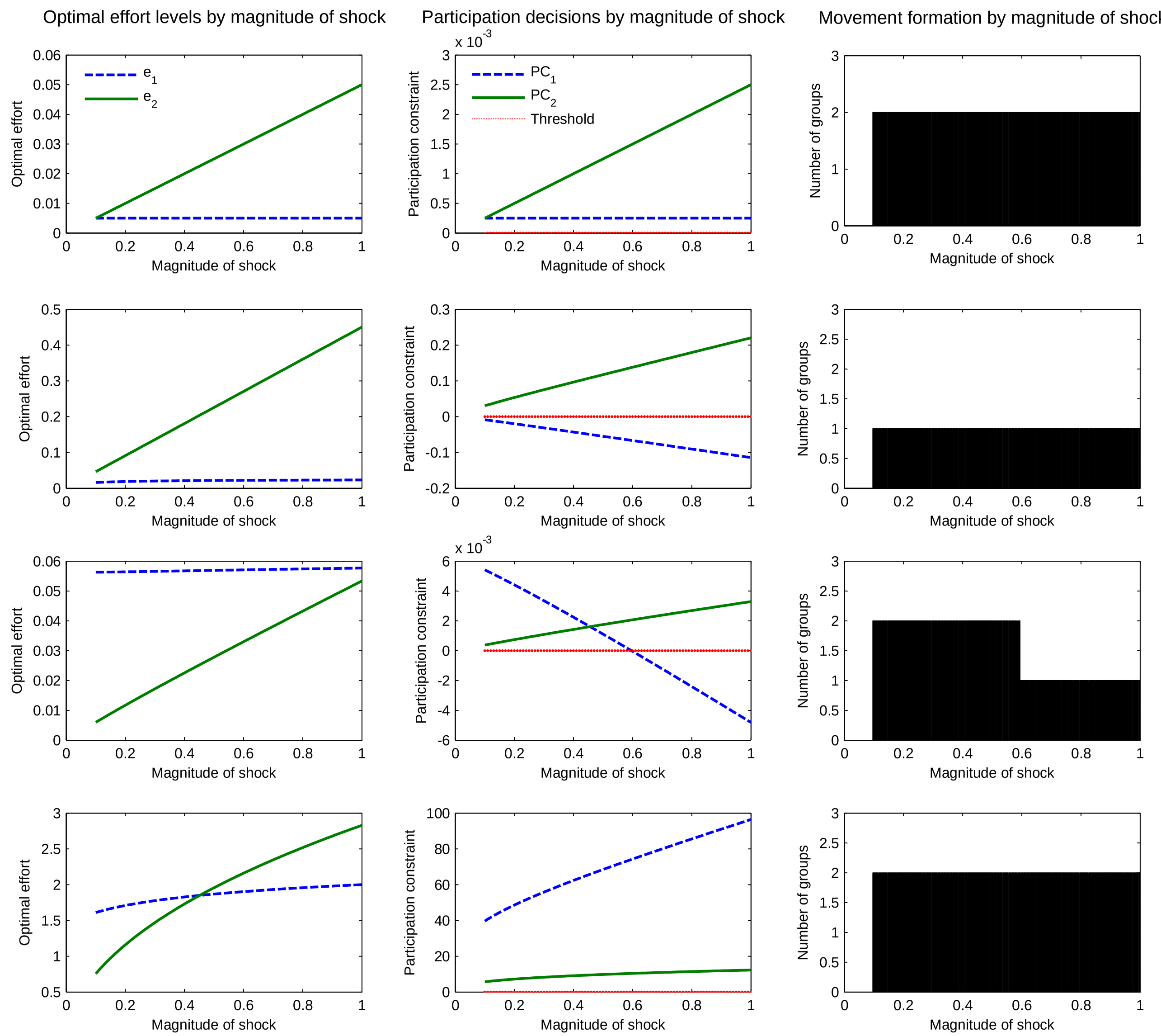

Figure 2: Effect of shocks on optimal effort levels, participation constraints, and social movement size.

spect to the two dimensions of the magnitude of the shock and the intensity of policy preferences, are illustrated in Figure 3 (a) and (b), respectively. As can be seen, the participation decision of the treatment becomes stronger as either the shock experienced or the intensity of policy preference (relative to the *status quo* of unity) increases, peaking when these are at the maximum values of the domains considered. The converse applies to the control: Its desire to participate in the movement is minimized just as the treatment's participation constraint attains its simulated maximum.

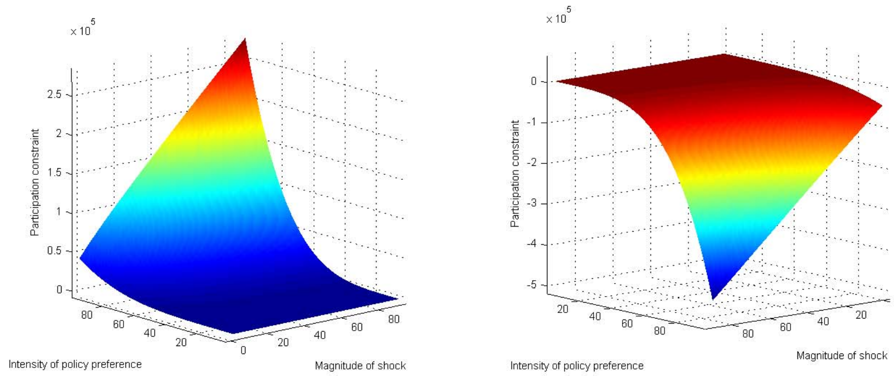

Figure 3: Participation decisions by magnitude of shock and intensity of policy preferences.

The case $J > 2$ presents situations of multiple equilibria that are sensitive to specific parameter assumptions, and hence is less insightful to analyze. Nevertheless, it is possible to forward some broad conjectures on the nature of group behavior in this case. To do so, we need to make some additional assumptions that restrict our set of possible equilibria. First, assume *weak asymmetry*; that is, economies are populated with $J - 1$ groups that are homogeneous, but distinct, from the jth group. Second, let homogeneous groups' choices reflect *movement momentum*, in the sense that their choice of participation is dependent only on their individual group's entry and exit, with none of the other groups leaving if they leave the movement. Since—by Example 1—symmetric agents will always form social movements with full participation, the set of equilibria can then be limited to the case examined above with just two groups. This allows social movements with a greater number of participating groups to form, and these display qualitative dynamics similar to Figure 2.

2.2 Group formation: Intra-group dynamics

While social movements are less likely to suffer from free-rider problems due to the non-negligible impact of their effort on realized policy and their different ideal policies, this is not the case for groups within the movement. The direct consequence of Assumption 2 is that group-specific preferred policy is a public good. Here, we relax this assumption by providing a framework for understanding intra-group dynamics that may lead to stable group policy preferences.[5]

Free-rider problems predicate an alternative modeling strategy for intra-group dynamics. More specifically, a system of endogenous enforcement needs to exist in order for groups to form, and be sustained. This underscores the importance to which the individuals that participate in the groups that constitute social movements are in fact embedded in their respective social structures and practices, as opposed to the movement *per se*. We therefore proceed by modeling a setup where individuals with bilateral relationships in groups have incentives not to free ride, due to an enforcement mechanism inherent in these relationships. In this regard, we draw on the literature on endogenous enforcement based on incentive slack (Bernheim & Whinston 1990; Bulow & Rogoff 1989; Fearon & Laitlin 1996).

Consider a group with three agents represented by the set $A = \{a, b, c\}$. These agents interact in bilateral relationships over an infinite horizon (or, alternatively, over a finite horizon with no known termination time). These relationships are summarized in Figure 4.

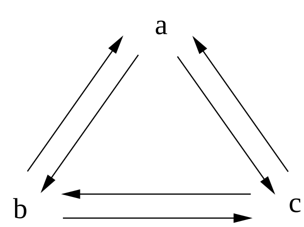

Figure 4: Agent relationships.

Assumption 4 (Limited enforceability)**.** For a given bilateral relationship between i and k, there will only be unidirectional incentive problems, such that at least one incentive compatibility constraint is slack. That is, either $v_{ik,0}(z,y) - v_{ik,0}(y,y) < \frac{\delta_i}{1-\delta_i}[v_{ik,0}(y,y) - v_{ik,0}(z,z)]$, or instead $v_{ki,0}(z,y) - v_{ki,0}(y,y) < \frac{\delta_k}{1-\delta_k}[v_{ki,0}(y,y) - v_{ki,0}(z,z)]$, but not both.

We employ this assumption primarily for tractability; the setup also limits the problem to situations where potentially interesting solutions may be found.

Definition 4. *Collectives* are bilateral relations where incentive compatibility constraints in the bilateral relationship between b and c, and between c and b, are slack. That is, $v_{bc,0}(z,y) - v_{bc,0}(y,y) < \frac{\delta_b}{1-\delta_b}[v_{bc,0}(y,y) - v_{bc,0}(z,z)]$ and $v_{cb,0}(z,y) - v_{cb,0}(y,y) < \frac{\delta_c}{1-\delta_c}[v_{cb,0}(y,y) - v_{cb,0}(z,z)]$.

[5]There are, clearly, other precedents in the study of group preference formation. There is a vast literature on non-cooperative group decisionmaking; the game-theoretic literature has focused primarily on bargaining situations, while the social choice literature has is based mainly on extensions of the Condorcet Jury Theorem. There is also a small literature on cooperative group decisionmaking; see Nehring (2004) for a recent example.

Assumption 5 (Collective formation)**.** If there is a bilateral relationship between b and c, it will be a collective.

This limits the problem to only the cases where there are potential enforceability problems between a and b and *vice versa*, as well as between a and c (and *vice versa*). We could envision this as either an institutional framework that guarantees enforceability in bilateral relations between b and c, or endogenously as a function of payoffs. For the latter case, this would be possible when we allow the payoffs in the game between b and c to differ from those between a and b, and a and c. In particular, this occurs when $v_{bc,0}(y, y)$ is sufficiently large relative to $v_{bc,0}(z, z)$, that is, $v_{bc,0}(y, y) \gg v_{bc,0}(z, z)$ (and *vice versa*). Alternatively, this assumption could hold even without $v_{bc,0}(y, y) \gg v_{bc,0}(z, z)$, so long as the discount factors δ_b and δ_c are sufficiently large.

Assumption 6 (Nash reversion strategy)**.** Agents employ a *Nash reversion strategy* in the group formation game for their interactions over the infinite horizon. More specifically, a player i plays along the path $Q(s)$ until someone defects, and the Nash equilibrium of the stage game $\{q_1^*, q_2^*\}$ is played thereafter.

This trigger strategy is fairly standard in the literature; it allows us to limit the scope of the game being played. Moreover, there is some limited experimental evidence that such a strategy is actually employed in repeated game settings, including that of the Prisoner's dilemma (Engle-Warnick & Slonim 2004; Selten & Stoecker 1986).

Definition 5 (Group formation equilibrium)**.** A (subgame perfect pure strategy) Nash equilibrium in the group formation game is given by a pair $\{q_1^*, q_2^*\}$ induced by the profile $s = (s_1, s_2) \in \Sigma_1 \times \Sigma_2$ such that $\forall i, k \in A, i \neq k$: $\left\{\nexists\, V'_{ik,0} \neq V^*_{ik,0} \text{ such that } V'_{ik,t} > V^*_{ik,t}\right\}$.

Proposition 2 (Group formation)**.** *For* $0 < \delta_i < 1$, $A = \{a, b, c\}$*, and relations of the form* $a \leftrightarrow b$ *and* $a \leftrightarrow c$*, for* $0 < \delta_i < 1$*, there exist Nash reversion strategies for sufficiently large transfers* $\tau > 0$ *such that group formation is sustainable, even if one incentive compatibility constraint is violated.*

Proof. See appendix. □

The proposition thus asserts that, with appropriate transfers, group formation in stage 0 of the game can be sustained. The proof relies on the existence of externality benefits accruing to an individual i for the maintenance of the relationship between two other individuals k and l. Since these benefits are idiosyncratic, groups are stable whenever this externality is sufficiently large. Individuals choose to remain in groups because of the transfer mechanism, which acts as a reward structure (or positive selective incentive). In exchange, they provide effort for the purposes of group activity. Group formation, therefore, is not inconsistent with the wider goals of a social movement. Additional intuition may be gained by the following example to Proposition 2.

Example 3. In a group formation equilibrium with the specific payoffs given above, and symmetric discount factors, group formation is sustainable if

$$\frac{\delta}{1-\delta}\left(\beta^{\epsilon}_{c,ab}+\beta_{ab}\right)\geq e_{ab}.$$

Proof. See appendix. □

The example implies that, with symmetric discount factors, a sufficient condition for sustainable group formation is when the next-period discounted value of the stream of both externality and actual benefits exceeds the value of effort. Individuals are willing to participate in groups since they receive not just the benefits of their own relationship with others in the group, but also from the existence of relationships between others in the group.

2.3 Social change equilibrium

While the underlying mechanisms for group and social movement formation are distinct, it is possible to embed these two processes into a single political-economic equilibrium that explains social change. To proceed, however, we need to restrict the extended game to accommodate the two dynamic processes that we have developed thus far. Effort enters as a disincentive to individual welfare, such that (2) is now $V_{ik,0}\left(q_{i,t},q_{k,t};\sigma_i\left(e_{ik}\right)\right)$, $\frac{\partial V_{ik,0}}{\partial\sigma_i}<0$, where $\sigma_i\left(e_{ik}\right)$ is a measure of the welfare costs of effort for individual i, and $k\in j$ is another agent within the group that i has a relationship with.

Assumption 7 (Group incentives)**.** For an individual $i\in j$: (a) At most one incentive compatibility constraint is violated; (b) i contributes an equal amount of effort to the group such that $e_i=\frac{e_j}{n_j}$, where $e_i=\sum_k^{n_j-1}e_{ik}$ is total individual effort.

This assumption ensures group existence, and sets the parameters for individual effort levels in a group. The following definition characterizes the equilibrium of the complete game.

Definition 6 (Social change equilibrium)**.** A (subgame perfect pure strategy) Nash equilibrium in the full-information social change game is a triplet $\left\{\left\{\mathbf{q}_i^*\right\}_{i\in A},\mathbf{e}^*,\mathbf{x}^*\right\}$ such that: (a) $\forall i,k\in A,i\neq k:\{\nexists V'_{ik,0}\neq V^*_{ik,0}$ such that $V'_{ik,0}>V^*_{ik,0}\ \forall t\}$; (b) $\forall j\in H\subseteq J:\{\nexists e'_j\neq e^*_j$ such that $U_j\left(e'_j\right)>U_j\left(e^*_j\right)\}$; (c) $\nexists\,\mathbf{x}'\neq\mathbf{x}^*$ such that $p\left(e\right)\sum_H\frac{1}{2}\left(x'-x_j\right)^2<p\left(e\right)\sum_H\frac{1}{2}\left(x^*-x_j\right)^2$.

This political equilibrium establishes the central result of our baseline model.

Proposition 3 (Social change)**.** *For a social movement where each group has an equal ability to influence policy through effort, a group j will participate in a social movement if there exists an individual $\tilde{i}\in j$ for which*

$$\beta^{\epsilon}_{\tilde{i},kl}\geq\hat{\tau}_{\tilde{i}k}+\sigma_k\left(e_k\mid\Pi\left(\mathbf{x}^*,e^*_j\right)\geq 0\right),\qquad(7)$$

with $\Pi\left(\mathbf{x}^{*}, e_{j}^{*}\right) \equiv \arg\max_{e_j}\{p\left(e\right) u_{j}\left[x_{j}^{*}\left(e\right), e_{j}\left(\xi_{j}\right)\right] + \left[1 - p\left(e\right)\right] u_{j}\left[\bar{x}_{j}, e_{j}\left(\xi_{j}\right)\right]\} =$
$p\left(\tilde{e}\right) u_{j}\left[x_{j}\left(\tilde{e}\right), 0\right] + \left[1 - p\left(\tilde{e}\right)\right] u_{j}\left[\bar{x}_{j}, 0\right]$ *and* $\tilde{i} \neq k, l$.

Proof. See appendix. □

What the proposition establishes is that if there exists an individual $\tilde{i}$ within a group j who benefits disproportionately from the continued maintenance of social relationships between others in the group, this individual can pay the majority to stay in the group and provide effort toward the social movement. The extent of this benefit, in turn, depends on the size of the group, its preferred policy position, the total group effort, and the size of the exogenous shock. This individual large player can also be thought of as a cohesive faction of moderately large players, such as a local middle class whose members all gain more from an anticipated social change due to improved business opportunities than do members of the poor majority, who can nonetheless be paid transfers to induce their participation in the groups that constitute the social movement.

3 Extensions

In this section, we explore several extensions to the baseline model developed in Section 2.

3.1 Social relationship mechanisms

Although we have only considered one form of social relationship mechanism in Subsection 2.3, it is possible to construe other mechanisms that could give rise to the social change equilibrium described in Definition 6.

One such mechanism is to consider the effect of linked games. Suppose agents in a community engage in pair-wise repeated games, the horizon is infinite or uncertain, information is perfect, and there are many identical agents. By the folk theorem any set of individually rational payoffs can be supported if the discount factor (corrected for uncertainty) is large enough. Suppose for concreteness that this game is a repeated Prisoners' Dilemma with mutual cooperation supported by grim trigger minimax strategies. Thus,

$$\gamma_i \leq \frac{\delta_i}{1 - \delta_i},$$

where $\gamma > 1 > 0 > -1$ are the payoffs. Suppose now a social movement opportunity arises, with optimal effort for the group as a whole given as before by e_j. To ensure the provision of $\frac{e_j}{n_j}$, agents could link the Prisoners' Dilemma game with the effort provision game such that either unilateral defection or effort provision below $\frac{e_j}{n_j}$ elicits the grim trigger punishment. Agents then cooperate in both games if $\gamma_i + \sigma_i\left(e_i \mid \Pi \geq 0\right) \leq \frac{\delta_i}{1-\delta_i}$.

An alternative approach is to consider the effect of punishment, or negative selective incentives. The single large player may also simply threaten other players to provide effort. Such a threat, of course, must be credible. Consider the case where there exists a subset $J' \subset J$ for whom the large player can maintain a credible threat. This game will then have a Nash equilibrium with $n_{j'} - 1 < n_j - 1$ players playing a repeated game with the large player, with defection incentives that are slack. The large player can then elicit from each of these i' players an effort level $e_{i'}$ satisfying $\gamma_{i'} + \sigma_{i'} (e_{i'} \mid \Pi \geq 0) \leq \frac{\delta_{i'}}{1-\delta_{i'}}$. It may be possible that, in total, the large player elicits $n_{i'} e_{i'}$ units of effort which may be greater than or equal to e_j.

These alternative social mechanisms can be relevant in specific real-world situations. The potential for linkages between issues secondary to the purposes of the movement appears to have been an important reason why participants in groups that comprise the anti-globalization movement have such diverse origins: Payoffs—external to any successful rejection of globalization—from membership in labor unions, neo-Nazi organizations, and environmentalist groups induce the support of individuals that would otherwise not hold strong views with respect to globalization *per se*, or would free-ride. Similarly, both black churches and colleges—examples of linkage-enforced social relationships—were prominent participants in the American civil rights movement.

3.2 Interactions with elites

As discussed in the introduction, elite complicity seems like a fairly strong condition to impose on social movement formation, and its empirical record has been less than stellar. Nonetheless, the involvement of an existing elite cadre does seem to be a relevant factor in explaining some of the dynamics of social movement formation. In discussing the movement that overthrew President Fujimori in Peru, it was "not the elite whom [Fujimori's secret-police chief] Montesinos feared, but the masses." (McMillan & Zoido 2004, p. 91). Nonetheless, the role of the elite-supported *El Comercio* and Channel N, while not "decisive... [was] potent" (McMillan & Zoido 2004, p. 90). Similarly, in the civil rights movement in the United States, disagreements among the existing elite led to support by some of these elites for social movement organizations (Jenkins & Eckert 1986), although this support was reactive, and in general not instrumental to the final outcome (McAdam 1999).

Elite behavior may be modeled in several ways. First, elites can respond to the same or different exogenous variables that groups in the movement respond to. For example, a financial crisis may impact the opportunity cost of resource deployment similarly for the social groups and the government by leading to a general economic crisis, or it may impact the government more if the government is relatively more dependent on the formal domestic or international financial sectors. Second, elite behavior changes may be derived as a response to group behavior changes. For example, since large shocks prompt social movements to form, elites can try to lower the mean or variance of the shock distribution $f(\xi)$. If the shock is the opportunity cost of time for movement participants, higher and more stable output in the economy, or less inequality, may work. Alternatively, elites can seek to resist social movement once formed, which could be modeled as a transformation of the success probability function to $\alpha p(e)$, $\alpha \in [0,1]$; further, $\alpha > 1$ could be used to reflect outside support or intervention in favor of the movement, or a weakening of the regime. Third, elites can induce a more favorable *status quo*,[6] thus lowering the *status quo* deviation, $\bar{x}$—and thereby possibly effectively buying off some groups. To the extent that this brings disutility for elites—since effort is costly and success uncertain for the movement—they may wish to only partially accommodate these groups. Relatedly, the regime can potentially bribe or collude with selected groups. Fourth, the regime may have, or accrue, special information or capabilities that make it costly for others to remove it; this is analogous to "poison pills" in the corporate governance literature. Similarly, as a social movement threatens to depose a still incumbent elite, it may be credible for the former to play "terminal period" strategies and gamble for resurrection, imposing large costs on the movement population (de Figueiredo, Jr. & Weingast 1999).

3.3 Asymmetric information

While the discussion in Subsection 2.1 allows for the failure of social movements to develop for particular configurations of preferences and shocks, one may wish to justify how movements may not arise even when agents are symmetric *in terms of preferences*. Essentially, we seek here to relax the strong implications of Example 1, by introducing a role for informational imperfections.

To do so, we allow, as before, the extent by which shocks impact a group to differ. However, these shocks are now private information, and so $\xi_j \in \Xi_j$, while still distributed with mean unity and support $(-\underline{\xi}, \overline{\xi})$, is now private information. Let Ξ_j contain a finite number of elements, such that we can define group j's beliefs about others' types as the conditional probability $r(\xi_{-j}|\xi_j)$. The Bayesian equilibrium of the game will now be such that a group j will participate in a social movement if and only if

$$\begin{aligned} &\arg\max_{e_j} \sum_{\xi_{-j}} \left\{ \begin{array}{c} p(e)\, r(\xi_{-j}|\xi_j)\, u_j \left[x_j^*(e), e_j(\xi_j, \xi_{-j})\right] \\ +\left[1-p(e)\right] r(\xi_{-j}|\xi_j)\, u_j \left[\bar{x}_j, e_j(\xi_j, \xi_{-j})\right] \end{array} \right\} \geq \\ &\sum_{\xi_{-j}} r(\xi_{-j}|\xi_j) \left\{ p(\tilde{e})\, u_j \left[x_j(\tilde{e}), 0\right] + \left[1-p(\tilde{e})\right] u_j \left[\bar{x}_j, 0\right] \right\}. \end{aligned} \tag{8}$$

This additional wrinkle is sufficient to break the strong symmetry results derived in Example 1. Denote the equilibrium effort level in the presence of asymmetric information by $e_j^*\left[r(\xi_{-j}|\xi_j)\right]$. In this case, the equilibrium response will once again be dependent on other groups' effort levels (or, more precisely, their expected effort levels given beliefs), and so the crucial simplifying result $x^* = \sum_H \frac{1}{h}\hat{x}_j$ will no longer hold in most cases, which is sufficient to overturn the results of the example. Thus, as was the case in the two-group numerical simulations, movements may or may not result, depending on the interaction

[6] This is another way to model redistribution; see also Acemoglu & Robinson (2000, 2001).

between the shocks experienced and the beliefs held by a group about other groups' shocks.

Asymmetric information may also enter if groups move sequentially in the participation stage (stage 1). For simplicity, let the critical level of the shock be $\hat{\xi}$, and let a group's participation decision be a function that maps $\Pi\left(\mathbf{x}^*, e_j^*\right)$ into $\{0, 1\}$, where 0 (1) indicates nonparticipation (participation). With perfect information, only groups that have realized $\xi_j > \hat{\xi}$ will choose participation. In this case, it may be possible to support beliefs such that a (perfect Bayesian) pooling equilibrium exists where both groups that do and do not exceed the critical $\hat{\xi}$ will choose to participate in the movement.

The sorts of informational imperfections that lead to social movement formation appears to have been an important reason why the disinformation campaign undertaken by Radio Télévision Libre de Mille Collines (RTLM) was so successful in mobilizing participants in the *Interahamwe* militia that perpetrated most of the Rwandan genocide (Prunier 1995). The radio propaganda campaign exploited a largely uneducated population that largely believed the content of the transmissions, which incited ordinary Hutus toward violence against Tutsis. These included racist smears (such as referring to Tutsis as *Inyenzi*, or cockroaches), but also inducements such as food, money, and the land of murdered Tutsis—inducements that would have lowered the expected cost of participation as well as beliefs about others' expected costs.

4 Historical Relevance

In this section we discuss the historical relevance of our model. In particular, we consider two case studies of social movements that illustrate the workings of some of the main mechanisms that have been put forward in the theoretical section: The 1997/98 Asian financial crisis in Indonesia, and ethnic conflict in the Congo over the period 1996–2001.

4.1 The Asian financial crisis and the Indonesian Revolution

The political transition that toppled the thirty-two-year presidency of General Haji Mohammad Soeharto can probably be traced to the 1996 leadership split between Megawati Sukarnoputri, the daughter of the previous president Soekarno, and Soeharto. Intra-regime friction had already been brewing in the previous year, and this was coupled with civilian dissent, mainly by student activist groups (Aspinall 1995). Some early demonstrations ensued, but the dissenting groups were generally unable to formally organize into a cohesive opposition force to the prevailing regime. However, the forced removal of Megawati as head of the *Partai Demokrasi Indonesia* led to mass demonstrations and later riots in Jakarta in July 1996. Despite this initial resistance, the *reformasi* movement—while signaling the potential for regime change—still failed to take off in a significant way, in part because the powerful Indonesian military, while demonstrating increasing dissatisfaction, nonetheless remained loyal to the administration.

It took the Asian financial crisis to coalesce additional opposition groups, such as nongovernmental organizations and the government-recognized Islamic organization *Muhammadiyah*, as well as proto-parties, into a credible political force. The crisis led to a plunge in the value of the rupiah; this, coupled with IMF-sponsored austerity measures, led to widespread increases in the prices of basic commodities such as rice, cooking oil, and fuel, as well as the removal of subsidies on many public services (a fall in the opportunity cost of effort or a deterioration of the status quo). The protests "gained force... in direct response to the collapse of the economy... and exploded... following the announcement of price increases on fuel and a number of other commodities." (Haggard 2000, p. 116).

In mid-May 1998, serious riots broke out throughout Jakarta, sparked by the killing of four students outside Triskati University by military forces. The riots left an estimated 1,200 dead, and major economic losses due to the damage and destruction of shophouses, residencies, and commercial buildings. Some reports suggest that the riots had an organized nature, and that they were possibly masterminded by elements within the military (van Klinken 1998). In any case, there participants in these riots were the urban poor (as opposed to student groups), who chose to be involved in the social violence more as opportunists who saw pecuniary and nonpecuniary benefits from such action. The combination of the riots and mass protests ushered in Soeharto's resignation on May 21.

The resignation of Soeharto, however, did not signal an end to the social conflict. By the second half of 1998, riots and social violence had become more widespread: The *reformasi* movement, previously concentrated in Jakarta and other urban centers, had now spread all over the country (Mathews 1998). The unifying call revolved around several basic themes: A reduction in the price of basic commodities, elimination of *korupsi*, *kolusi*, and *nepotisme* (corruption, collusion, and nepotism), and political reform. Unfortunately, while these movements often started off as reactions to the abuses of the old regime, they also unleashed forces of a more malevolent nature: Sectarian conflict, violent crime, and armed secessionism. The most severe among these was the communal violence in Maluku: The crisis had spurred Christian-Muslim clashes throughout the former Dutch East Indies. By the cessation of the worse of the fighting in June 2000, an excess of 1,500 were estimated dead.

The contentious politics that characterized the Indonesian Revolution of 1998 provides support for the processes described in our theoretical model. First, Aspinall (2005) documents the diversity of strategies and techniques employed by various opposition groups to bring about political change; opposition groups engaged in processes that reflected definite strategic behavior, with groups undertaking political action based on fears of losing ground to rival groups. The *reformasi* movement, therefore, was as much a competitive enterprise as it was a cooperative one. Moreover, elite involvement in overturning the regime was, at best, ambivalent and acquiescent, rather than active. For example, the turning point of the political crisis that ultimately led to Soeharto's resignation was marked more by the desertion by this elite, rather than their support of the movement *per se*.

Second, the economic shock that resulted from the Asian financial crisis very likely lowered the opportunity cost of entry for a much wider number a opposition groups—the entry of which precipitated the spread of the protests beyond urban centers. While we do not have the counterfactual, it is doubtful that the regime—which had survived previous threats to its legitimacy, usually by force—would have collapsed in the absence of concomitant confrontations all over the archipelago.

Third, other elements of our theoretical model appear to be important. For example, organizational leadership was clearly necessary (and evident) in ensuring that the movement's disparate participants were united in their ultimate endeavor. Likewise, informational imperfections may have contributed to the participation of groups that may otherwise have not been involved. Duncan (2005), for example, discusses how competing perceptions of the clashes between Christians and Muslims in Maluku may have played a role in fermenting and abetting a continuance of the sectarian violence.

4.2 Uprisings against the Mobutu and Kabila presidencies in the Congo

Civil conflict and violence against civilians has been endemic in Central Africa since independence, and the most serious recent conflict—centered in the Eastern part of the Democratic Republic of the Congo (DRC, formerly Zaire)—has involved up to nine countries. Many observers term it Africa's World War. This conflict began in May 1998, fifteen months after a successful rebellion in the DRC led to the overthrow of President Mobutu and the assumption of power by rebel leader Laurent Kabila. An initial peace agreement was brokered in July 1999, with further withdrawal of Rwandan troops from the Eastern parts of the country in 2002. Two months later, the remaining parties in the conflict signed an accord to end fighting and establish a transitional government in July 2003. Nonetheless, fighting has continued, and the disarmament, demobilization, and resettlement of displaced people have yet to be accomplished. Large parts of the country are not under government control, and large foreign troop contingents remain present.

The success of Laurent Kabila in overthrowing Mobutu, who had ruled since a coup in 1965, reflected active support from the governments of Rwanda and Uganda. This was enabled and motivated in turn by the coming to power in Rwanda in July of 1994 of the ethnic Tutsi-dominated Rwandese Patriotic Front (RPF). Tutsis living in Uganda and the DRC made up a significant part of the fighting forces of Kabila. Indeed, the Hutu and Tutsi populations in the DRC, despite living in the country since independence, had long been subjected to violence and discrimination, being considered non-Congolese by the local Hunde and Nyanga ethnic groups (Minorities at Risk Project 2006) and targeted by local governments.

The alliance between Hutu and Tutsi in the DRC was split by the arrival in July to August 1994 of about 720,000 Hutus, including many from the defeated *Interahamwe* militias that had just carried out the Rwandan genocide. In October 1996, the deputy governor of South Kivu, Lwasi Ngabo Lwabanji, told ethnic Tutsis to leave the Congo—with potential threats of ethnic cleansing—and this sparked a Congolese Tutsi rebellion. This is turn encouraged opposition groups, separatists and pro-democracy activists throughout the Congo. The rebel forces became known as the *Alliance des Forces Démocratiques pour la Libération du Congo-Zaïre* (Minorities at Risk Project 2006).

However, political protests and ethnic massacres against—and perpetrated by—Congolese Tutsis had been ongoing since March 1993. In September 1996, the Rwandan government, concerned with the fate of the Tutsis and eager to curtail the security threat from remobilizing Hutu militants in exile supported by Mobutu, sent troops across the border to protect and support the Congolese Tutsis (Olsson & Fors 2004). Uganda, also housing a Tutsi population and concerned with the fate of Tutsis in Zaire and the survival of the new Rwandan government (Clark 2001), sent troops supporting the rebellion as well. Once underway, the rebellion was joined by other anti-Mobutu groups, including the faction led by Kabila. Mobutu's disorganized and poorly motivated army offered slight resistance and Kinshasa surrendered in May 1997.

Aspects of Kabila's rebellion in 1996 are well captured by our theoretical model. The resident Tutsis in the DRC certainly were unhappy with their *status quo* under Mobutu, suffering violence, discrimination, and lack of civil rights. This, however, they shared with several other ethnic groups in Zaire, including Congolese Hutus. But the genocide and warfare in Rwanda, coupled with the massive inflow of mainly Hutu refugees (including many of the anti-Tutsi *Interahamwe* militia), may have acted as an exogenous shock that changed the cost-benefit ratio of rebellion: The cost of doing nothing rose with the growing presence of Hutu militants, and rapidly increasing Congolese Hutu-Congolese Tutsi tensions. This was abetted by increasing anti-Tutsi sentiment among the local authorities and other ethnic groups in Zaire, and international support from Rwanda and Uganda. Whether Kabila was a particularly able leader is unclear, and although he treated Tutsis unfavorably after assuming the presidency, he may have seemed a promising leader *ex ante*. Once begun, the Congolese Tutsi rebellion prompted participation by non-Tutsi groups in the Congo, and groups from Rwanda and Uganda, possibly by a similar mechanism.

Similarly, aspects of the insurrection against Kabila in 1998—backed by governments in Rwanda and Uganda in a turnaround of their erstwhile support for Kabila—may be explained by our model as well. First, the timing appears consistent with Rwanda and Uganda having learned enough to conclude that Kabila had not been a good bet after all: Beyond a corrupt and incompetent rule, he dismissed Tutsis from top military and political positions, demanded the withdrawal of Rwandan troops, and failed to protect Congolese Tutsis from ever more serious violence from other groups (Minorities at Risk Project 2006). This changed information structure resulted in Rwanda and Uganda supporting the *Mouvement de Libération du Congo* (MLC).

Second, the decision of Angola, Chad, Namibia, Sudan, and Zimbabwe to intervene in support of Kabila is also consistent with our theory. These nations, in spite of their apparent heterogeneity, face a sufficiently similar objective: Claim to the immense mineral wealth of the DRC, including diamonds, gold, and coltan (United Nations 2001a,b). Common security motives may also have been important for some countries; for example, the Angolan government sought to erode União Nacional para a Independência Total de Angola (UNITA) rebel bases in the Congo in order to guarantee its border (Olsson & Fors 2004). Participants in the insurrection were definitely purposeful, strategic actors.

5 Conclusion

The factors that give rise to social movements and political change are varied and complex, as testified to by the vast literature analyzing these phenomena across many social scientific disciplines. Nonetheless, our argument is that such movements can be understood essentially as a combination of both intra- and inter-group dynamics, where groups overcome traditional free-rider problems by a combination of both social relationship mechanisms as well as strategic behavior depending on, *inter alia*, the extent to which groups experience exogenous shocks to their opportunity cost structures. We believe that our work provides an important insight into the dynamics of real-world social movements: Movements need not be critically dependent on interactions between elites versus the populace, but can in fact arise endogenously from changes in participants' local environments. Examples of such changes are the occurrence of financial crises and political changes in neighboring countries.

Understanding the causal importance of these mechanisms may also aid policymaking, especially with regard to ushering in democratic transition in autocratic and semi-autocratic countries. To the extent that such transitions are viewed as desirable, our paper suggests that these may be fostered by amplifying channels where exogenous shocks may have a greater impact, essentially helping people help themselves. Of course, social movement activity often entails significant costs for participants and non-participants alike (for example, in the case of revolutionary movements), and these must be considered as well. The role of advances in information and communications technology, such as the Internet, in lowering the costs of coordination may be important in this regard. However, the ubiquity of such technologies may also have a dark side: Cellphone text-messaging has been implicated as the basis for organizing the racially-motivated riots in Sydney in late 2005.

The primary shortcoming of the present paper is the absence of formal empirical tests of our main propositions. More specifically, by assuming functional forms similar to the one used for our examples, it may be possible to test the structural equation that undergirds Propositions 1 and 2. The household- and community/facility-level data from the Indonesian IFLS dataset, circulated by the RAND Corporation, is ideal for this purpose. Such an empirical investigation could proceed by estimating a two-stage least squares system of two simultaneous equations, corresponding to the inter- and intra-group formation equations in our theoretical model. The former equation would seek to capture the determinants of group participation in economic activism, by region, while controlling for important region-level factors such as regional socioeconomic development. The latter would regress local community participation on household socioeconomic variables as well as proxies for the strength of social relationship effects, such as village size and the presence of community institutions. We plan to undertake such testing in future work.

References

Acemoglu, K. Daron & James A. Robinson (2000). "Why Did the West Extend the Franchise? Democracy, Inequality, and Growth in Historical Perspective". *Quarterly Journal of Economics* 115(4) (November): 1167–1199

Acemoglu, K. Daron & James A. Robinson (2001). "A Theory of Political Transitions". *American Economic Review* 91(4) (September): 838–863

Aspinall, Edward (1995). "Students and the Military: Regime Friction and Civilian Dissent in the Late Suharto Period". *Indonesia* 59 (April): 21–44

Aspinall, Edward (2005). *Opposing Suharto: Compromise, Resistance, and Regime Change in Indonesia*. Stanford University Press

Bernheim, B. Douglas & Michael D. Whinston (1990). "Multimarket Contact and Collusive Behavior". *RAND Journal of Economics* 21(1) (Spring): 1–26

Breton, Albert & Raymond Breton (1969). "An Economic Theory of Social Movements". *American Economic Review* 59(2) (May): 198–205

Bulow, Jeremy I. & Kenneth S. Rogoff (1989). "A Constant Recontracting Model of Sovereign Debt". *Journal of Political Economy* 97(1) (February): 155–178

Chwe, Michael Suk-Young (1999). "Structure and Strategy in Collective Action". *American Journal of Sociology* 105(1) (July): 128–156

Clark, John F. (2001). "Explaining Ugandan Intervention in Congo: Evidence and Interpretations". *Journal of Modern African Studies* 39(2) (June): 261–287

Darnovsky, Marcy, Barbara L. Epstein & Richard Flacks (editors) (1995). *Cultural Politics and Social Movements*. Philadelphia, PA: Temple University Press

de Figueiredo, Jr., Rui J. P. & Barry R. Weingast (1999). "The Rationality of Fear: Political Opportunism and Ethnic Conflict". In Barbara F. Walter & Jack L. Snyder (editors), *Civil Wars, Insecurity and Intervention*, pp. 261–302. New York, NY: Columbia University Press

Drazen, Allan & Vittorio Grilli (1993). "The Benefit of Crises for Economic Reforms". *American Economic Review* 83(3) (June): 598–607

Duncan, Christopher R. (2005). "The Other Maluku: Chronologies of Conflict in North Maluku". *Indonesia* 80 (October): 53–80

Engle-Warnick, Jim & Robert L. Slonim (2004). "The Evolution of Strategies in a Repeated Trust Game". *Journal of Economic Behavior & Organization* 55(4) (December): 553–573

Fearon, James D. & David D. Laitlin (1996). "Explaining Interethnic Cooperation". *American Political Science Review* 90(4) (December): 715–35

Fisher, William F. (1997). “Doing Good? The Politics and Antipolitics of NGO Practices”. *Annual Review of Anthropology* 26: 439–464

Frohlich, Norman & Joe A. Oppenheimer (1970). “I Get by with a Little Help from My Friends”. *World Politics* 23(1) (October): 104–120

Frohlich, Norman, Joe A. Oppenheimer & Oran R. Young (1971). *Political Leadership and Collective Goods*. Princeton, NJ: Princeton University Press

Fudenberg, Drew & Eric S. Maskin (1986). “The Folk Theorem in Repeated Games with Discounting or with Incomplete Information”. *Econometrica* 54(3) (May): 533–554

Granovetter, Mark (1978). “Threshold Models of Collective Behavior”. *American Journal of Sociology* 83(6) (May): 1420–1443

Grossman, Gene M. & Elhanan Helpman (1994). “Protection for Sale”. *American Economic Review* 84(4) (September): 833–850

Grossman, Herschel I. (1991). “A General Equilibrium Theory of Insurrections”. *American Economic Review* 84(3) (September): 912–921

Haggard, Stephan (2000). *The Political Economy of the Asian Financial Crisis*. Washington, DC: Institute for International Economics

Hulme, David & Michael Edwards (editors) (1997). *NGOs, States, and Donors: Too Close for Comfort*. New York, NY: St Martin’s Press

Jackson, Matthew O. (2005). “A Survey of Models of Network Formation: Stability and Efficiency”. In Gabrielle Demange & Myrna H. Wooders (editors), *Group Formation in Economics: Networks, Clubs and Coalitions*. Cambridge, England: Cambridge University Press

Jenkins, J. Craig & Craig M. Eckert (1986). “Channeling Black Insurgency: Elite Patronage and Professional Social Movement Organizations in the Development of Black Movement”. *American Sociological Review* 51(6) (December): 812–829

Kim, Hyojoung & Peter S. Bearman (1997). “The Structure and Dynamics of Movement Participation”. *American Sociological Review* 62(1) (February): 70–93

Labán, Raúl & Federico Sturzenegger (1994). “Fiscal Conservatism as a Response to a Debt Crisis”. *Journal of Development Economics* 45(2) (December): 305–324

Lehmann, David (1990). *Democracy and Development in Latin America: Economics, Politics, and Religion in the Postwar Period*. London, England: Polity Press

Marwell, Gerald & Pamela E. Oliver (1993). *The Critical Mass in Collective Action: A Micro-Social Theory*. New York, NY: Cambridge University Press

Mathews, Richard (1998). “The Impact of the May Crisis in Bali, West Nusa Tenggara, East Nusa Tenggara and East Timor”. In Geoff Forrester & Ronald J. May (editors), *The Fall of Soeharto*, chapter 3, pp. 70–92. Bathurst, NSW: Crawford House Publishing

McAdam, Doug (1999). *Political Process and the Development of Black Insurgency, 1930–1970*. Chicago, IL: University of Chicago Press, second edition

McAdam, Doug, John D. McCarthy & Mayer N. Zald (1996). “Introduction: Opportunities, Mobilizing Structures, and Framing Processes—Towarda Synthetic, Comparative Perspective on Social Movements”. In Doug McAdam, John D. McCarthy & Mayer N. Zald (editors), *Comparative Perspectives on Social Movements: Political Opportunities, Mobilizing Structures, and Cultural Framings*. Cambridge, England: Cambridge University Press

McCarthy, John D. & Mayer N. Zald (1977). "Resource Mobilization and Social Movements: A Partial Theory". *American Journal of Sociology* 82(6) (November): 1212–1241

McMillan, John & Pablo Zoido (2004). "How to Subvert Democracy: Montesinos in Peru". *Journal of Economic Perspectives* 18(4) (Fall): 69–92

Minorities at Risk Project (2006). *Minority Group Chronologies*. College Park, MD: University of Maryland. URL `http://www.cidcm.umd.edu/inscr/mar/`

Murphy, Kevin M. & Andrei Shleifer (2004). "Persuasion in Politics". *American Economic Review* 94(2) (May): 435–439

Nehring, Klaus (2004). "The Veil of Public Ignorance". *Journal of Economic Theory* 119(2) (December): 247–270

Oliver, Pamela E. (1993). "Formal Models of Collective Action". *Annual Review of Sociology* 19: 271–300

Oliver, Pamela E. & Gerald Marwell (1992). "Mobilizing Technologies for Collective Action". In Aldon D. Morris & Carol McClurg Mueller (editors), *Frontiers in Social Movement Theory*, pp. 251–271. New Haven, CT: Yale University Press

Olson, Mancur (1971). *The Logic of Collective Action: Public Goods and the Theory of Groups*. Harvard Economic Studies. Cambridge, MA: Harvard University Press, revised edition

Olsson, Ola & Heather Congden Fors (2004). "Congo: The Prize of Predation". *Journal of Peace Research* 41(3) (May): 321–336

Pommerenke, Kai (2005). *Cooperation with Rivals*. Ph.D. thesis, University of California, Santa Cruz

Prunier, Gérard (1995). *The Rwanda Crisis: History of a Genocide*. New York, NY: Columbia University Press

Selten, Reinhard & Rolf Stoecker (1986). "End Behavior in Sequences of Finite Prisoner's Dilemma Supergames". *Journal of Economic Behavior & Organization* 7(1) (March): 47–70

Srbljinovic, Armano, Drazen Penzar, Petra Rodik & Kruno Kardov (2003). "An Agent-Based Model of Ethnic Mobilization". *Journal of Artificial Societies and Social Simulation* 6(1) (January): Article 1

Tarrow, Sidney (1998). *Power in Movement: Social Movements and Contentious Politics*. Cambridge Studies in Comparative Politics. Cambridge, England: Cambridge University Press, second edition

United Nations (2001a). "Addendum to the Report of the Panel of Experts on the Illegal Exploitation of Natural Resources and Other Forms of Wealth of the Democratic Republic of the Congo". *Technical Report S/2001/1072*, United Nations, New York, NY

United Nations (2001b). "Report of the Panel of Experts on the Illegal Exploitation of Natural Resources and Other Forms of Wealth of the Democratic Republic of the Congo". *Technical Report S/2001/357*, United Nations, New York, NY

van Klinken, Gerry (1998). "The May Riots". *Inside Indonesia* 63: 1–3

Appendix

A.1 Proofs

Proof of Proposition 1. To demonstrate necessity, we solve the game by backward induction. In stage 3, the entrepreneur will take total effort e as given, and solves the program given by (3). This is then

$$\max_{\mathbf{x}} \sum_{j=1}^{H} \omega_j \cdot Eu_j \left[x_j \left(e\right), e_j \left(\xi_j\right)\right].$$

The first order necessary condition is

$$\sum\nolimits_{H} \omega_j \frac{\partial Eu_j \left[x_j \left(e\right), e_j \left(\xi_j\right)\right]}{\partial x_m} = 0 \quad \forall m \in M.$$

This then yields, for each policy dimension m, the implied policy vector $x^* \left(e\right)$, which we collect into an $\left(m \times 1\right)$ vector $\mathbf{x}^* \left(e\right)$. Denote $x_j^* \left(e\right) = \|\mathbf{x}^* \left(e\right) - \hat{\mathbf{x}}_j\|$. In stage 2, groups choose their effort levels, taking into account the chosen policy. Each group j will solve the problem in (5). This is then

$$\max_{e_j} \left\{p\left(e\right) u_j \left[x_j^* \left(e\right), e_j \left(\xi_j\right)\right] + \left[1 - p\left(e\right)\right] u_j \left[\bar{x}_j, e_j \left(\xi_j\right)\right]\right\},$$

which yields an optimal effort level for each group, e_j^*. Finally, in stage 1, groups take into account the above result and make their participation decisions according to (4). Simple substitution yields the result of the proposition. To see sufficiency, suppose that instead

$$\begin{aligned} &\arg\max_{e_j} \left\{p\left(e\right) u_j \left[x_j^* \left(e\right), e_j \left(\xi_j\right)\right] + \left[1 - p\left(e\right)\right] u_j \left[\bar{x}_j, e_j \left(\xi_j\right)\right]\right\} \leq \\ &p\left(\tilde{e}\right) u_j \left[x_j \left(\tilde{e}\right), 0\right] + \left[1 - p\left(\tilde{e}\right)\right] u_j \left[\bar{x}_j, 0\right], \end{aligned}$$

is an equilibrium. However, this is a contradiction of the definition of the social movement equilibrium given in Definition 3, which concludes the proof. □

Proof of Example 1. We again solve by backward induction. In stage 3, the entrepreneur will take effort total effort e as given, and solves the program given by (3). This is equivalent to

$$\min_{\mathbf{x}} \sum_{j=1}^{H} \omega_j \cdot \frac{1}{2} \left(\|\mathbf{x}\left(e\right) - \hat{\mathbf{x}}_j\|\right)^2.$$

The first order condition is

$$\sum\nolimits_{H} \omega_j \left(\mathbf{x} - \hat{\mathbf{x}}_j\right) = 0 \quad \forall m \in M.$$

This yields, for each policy dimension m, the result $x^* = \sum_{H} \omega_j \hat{x}_j$, which we collect into $\mathbf{x}^* \left(e\right)$. Now, denote $x_j^* \left(e\right) = \|\mathbf{x}^* \left(e\right) - \hat{\mathbf{x}}_j\|$. In stage 2, each group

j chooses its effort level in the problem in (5):

$$\max_{e_j} \left\{ p\left(e\right) \left[-\frac{1}{2} x_j^*\left(e\right)^2 \right] + \left[1 - p\left(e\right)\right] \left[-\frac{1}{2} \bar{x}_j^2 \right] - \frac{1}{2\xi_j} e_j^2 \right\}.$$

Let the probability $p(e) = \frac{e}{\bar{e}}$, where $\bar{e} = \sum_H \bar{e}_j$, where $\bar{e}_j$ is the minimal effort each group would need to exert in order for a given movement to succeed. Substituting this into the above, and solving, yields

$$e_j = \frac{\xi_j}{2\bar{e}} \left[x_j^*\left(e\right)^2 + \bar{x}_j^2 \right] \quad \forall j \in H. \tag{A.1}$$

Note that the optimal effort in this case is defined implicitly, since $x_j^*\left(e\right)$ remains a function of e_j. If we denote e_j^* as the optimal effort after isolating e_j in the above equation, we obtain from the participation condition the critical level of the shock:

$$\xi_j \geq \frac{e_j^{*2}}{\frac{\tilde{e}}{\bar{e}} \left[x_j\left(\tilde{e}\right)^2 - \bar{x}_j^2 \right] - \frac{e^*}{\bar{e}} \left[x_j^*\left(e\right)^2 - \bar{x}_j^2 \right]}, \tag{A.2}$$

where we have used the fact that $\tilde{e}_j = 0$. For the case of symmetric groups, the implemented policy now weights each group's ideal policy preference intensity equally, such that $x_j^* = \|\mathbf{x}^* - \hat{\mathbf{x}}_j\|$, where $x^* = \sum_H \frac{1}{h} \hat{x}_j$. Making the necessary substitutions for x_j^* and e_j^* and simplifying gives us

$$(h - g) \geq \frac{1}{2}, \tag{A.3}$$

where we have exploited the symmetry of groups such that $x_j^*\left(e\right) = x_j\left(\tilde{e}\right) = 0$. Since $h > g$, the inequality will always satisfied, which establishes the proposition as claimed. □

Proof of Example 2. Using the equation for optimal effort given by (A.1), it is straightforward to demonstrate the following by application of the implicit function theorem:

$$\frac{\partial e_j^*}{\partial \xi_j} = \frac{x_j^{*2} + \bar{x}_j^2}{2\bar{e}\left[1 - \frac{\xi_j x_j^*}{\bar{e}} \cdot \frac{\partial x_j^*}{\partial e_j}\right]} > 0 \tag{A.4a}$$

$$\frac{\partial e_j^*}{\partial \bar{x}_j} = \frac{2\xi_j \bar{x}_j}{2\bar{e}\left[1 - \frac{\xi_j x_j^*}{\bar{e}} \cdot \frac{\partial x_j^*}{\partial e_j}\right]} > 0 \tag{A.4b}$$

$$\frac{\partial e_j}{\partial \bar{e}} = \frac{-\xi_j\left(x_j^{*2} + \bar{x}_j^2\right)}{2\bar{e}^2\left[1 - \frac{\xi_j x_j^*}{\bar{e}} \cdot \frac{\partial x_j^*}{\partial e_j}\right]} < 0 \tag{A.4c}$$

where, to sign (A.4a)–(A.4c), we have used the fact that $\frac{\partial x_j^*}{\partial e_j} = -\frac{x_j^*}{e} < 0$, as well as the envelope theorem for (A.4a) and (A.4b), which guarantees that effort levels e_{-j} for a group $-j \neq j$ are such that $\frac{\partial e_{-j}}{\partial \xi_j} = \frac{\partial e_{-j}}{\partial e_j} \cdot \frac{\partial e_j}{\partial \xi_j} \approx 0$ and $\frac{\partial e_{-j}}{\partial \bar{x}_j} = \frac{\partial e_{-j}}{\partial e_j} \cdot \frac{\partial e_j}{\partial \bar{x}_j} \approx 0$. Note that the inability to sign the change in *equilibrium* effort levels with respect to $\bar{e}$ stems from the fact that a change in $\bar{e}$ affects all groups directly, which limits the application of the envelope theorem to (A.4c). □

Proof of Proposition 2. Before we proceed with the proof, it is useful to establish a lemma for the case where the discount rate for players in a bilateral relationship is large enough to sustain cooperation.

Lemma 1 (Folk theorem)**.** *For $\delta_i > \bar{\delta}_i > 0$, a cooperative outcome path $Q = \{q_{i,t}, q_{k,t}\}_{t=1}^{\infty} = \{y, y\}_{t=1}^{\infty}$ can be sustained with a Nash reversion strategy, where $\bar{\delta}$ is the highest discount factor that can sustain cooperation, given the payoffs $V_{ik,0}$.*

Proof. The lemma is a direct result of the Folk Theorem for repeated games applied to the context here; see Fudenberg & Maskin (1986) for a formal treatment of the problem. □

By Lemma 1, there will always exist cooperative outcomes with Nash reversion strategies when $\delta_i > \bar{\delta}_i$ for both players in a bilateral relationship. Together with Assumption 4, this necessary (though not sufficient) condition suggests that we only need to consider cases where one (and only one) incentive compatibility constraint is violated in a bilateral relationship. We therefore limit ourselves to the analysis of subcases with unidirectional incentive problems in the triad $a \leftrightarrow b$, $a \leftrightarrow c$. These are summarized below.

Subcase	Interactions	Subcase	Interactions
(i)	$a \rightarrow b, c \rightarrow a$	(iii)	$b \rightarrow a, a \rightarrow c$
(ii)	$a \rightarrow b, a \rightarrow c$	(iv)	$b \rightarrow a, c \rightarrow a$

We consider these sub-cases in turn. For any of these sub-cases, if both incentive compatibility constraints are violated, then by Lemma 1, group formation is not sustainable. Therefore, for the following, we will consider only cases with one constraint being violated. (i) Without loss of generality, let $v_{ab,0}(z,y) - v_{ab,0}(y,y) > \frac{\delta_a}{1-\delta_a}[v_{ab,0}(y,y) - v_{ab,0}(z,z)]$ as well as $v_{ca,0}(z,y) - v_{ca,0}(y,y) < \frac{\delta_c}{1-\delta_c}[v_{ca,0}(y,y) - v_{ca,0}(z,z)]$. Let the Nash reversion strategy for c be $s_{c,t}$: {"Make transfer τ to $a \Leftrightarrow a$ has cooperated against b in all past periods, otherwise play the single-period Nash strategy"}. Then there can potentially exist values of the transfer $0 < \tau < \hat{\tau}$ such that the direction of the inequality in the first incentive compatibility constraint reverses, while the inequality in the latter constraint remains unchanged. The new incentive constraints of relevance are then $v_{ab,0}(z,y) - v_{ab,0}(y,y) \leq \frac{\delta_a}{1-\delta_a}[v_{ab,0}(y,y) - v_{ab,0}(z,z) + \tau_{ca}]$ and $v_{ca,0}(z,y) - v_{ca,0}(y,y) \leq \frac{\delta_c}{1-\delta_c}[v_{ca,0}(y,y) - v_{ca,0}(z,z) - \tau_{ca}]$. $\hat{\tau}_{ca}$ is then the maximum transfer that can be effected while preserving the inequality in the second constraint. Note that the existence of such a transfer is not guaranteed. Specifically, the upper bounds of $a \rightarrow b$ cooperation is determined by the *externality* benefits $\beta^{\epsilon}_{c,ab}$ accruing to c from the maintenance of a relationship between a and b. Formally, such a transfer exists if and only if $\frac{\hat{\tau}_{ca}}{1-\delta_c} \leq \frac{\beta^{\epsilon}_{c,ab}}{1-\delta_c} \Leftrightarrow \hat{\tau}_{ca} \leq \beta^{\epsilon}_{c,ab}$. (ii) Without loss of generality, let $v_{ab,0}(z,y) - v_{ab,0}(y,y) > \frac{\delta_a}{1-\delta_a}[v_{ab,0}(y,y) - v_{ab,0}(z,z)]$ and also let $v_{ac,0}(z,y) - v_{ac,0}(y,y) > \frac{\delta_a}{1-\delta_a}[v_{ac,0}(y,y) - v_{ac,0}(z,z)]$. By Assumption 4, we also have $v_{ca,0}(z,y) - v_{ca,0}(y,y) < \frac{\delta_c}{1-\delta_c}[v_{ca,0}(y,y) - v_{ca,0}(z,z)]$. The rest of the proof then follows that for (i). (iii) Analogous to (i). (iv) Analogous to (ii).

Remark. It may be useful to view Proposition 2 as a form of a dynamic Coase theorem for ongoing relationships where trigger strategies need be applied because the conventional contract enforcement otherwise required for the Coase theorem to work does not work here. The externality embodied in β^{ϵ} is internalized by b because of the credible threat by c to cease transfers τ should there be noncooperation between a and b.

□

Proof of Example 3. Absent of transfers, the incentive constraints are given by $e_{ab} > \frac{\delta_a}{1-\delta_a}\beta_{ab}$ and $e_{ca} < \frac{\delta_c}{1-\delta_c}\beta_{ca}$, which, with transfers from c to a, become $e_{ca} \leq \frac{\delta_c}{1-\delta_c}(\beta_{ca} - \tau_{ca})$ and $e_{ab} \leq \frac{\delta_a}{1-\delta_a}(\beta_{ab} + \tau_{ca})$. The externality benefits must be such that $\frac{\hat{\tau}_{ca}}{1-\delta_c} \leq \frac{\beta^{\epsilon}_{c,ab}}{1-\delta_c}$. Manipulation of these last two equations gives the system

$$\frac{\tau_{ca}}{1-\delta_a} \geq \frac{e_{ab}}{\delta_a} - \frac{\beta_{ab}}{1-\delta_a},$$
$$\frac{\tau_{ca}}{1-\delta_c} \leq \frac{\beta^{\epsilon}_{c,ab}}{1-\delta_c},$$

which, by imposing $\delta_a = \delta_c$, yields the result in the example. □

Proof of Proposition 3. By Proposition 2, group formation is stable when

$$\beta_{i,kl} \geq \tau_{ik} + \sigma_k(e_k).$$

Moreover, by Assumption 7(a) and (b), the total amount of effort provided by individual k will be

$$e^*_k = \frac{e^*_j}{n_j}$$

Together with (6), successive substitutions into the equations above yields the result as shown. □

A.2 Simulation details

Simulations were performed in MATLAB. The optimal levels of effort were found using Broyden's rootfinding algorithm, with initial values of 0.2. For ease of interpretation, the values for the participation constraint were transformed such that participation occurs when the calculated value is positive. The results were not sensitive to the magnitude of parameter values chosen, although axes displayed appropriate re-scaling. For the two-group case, the subgame perfect pure Nash equilibrium defined in Definition 3 is well defined: Note that by Example 1, there will always be at least one group that chooses to enter the social movement (since with a single group (A.3) will be $(1-0) \geq \frac{1}{2}$), and so the calculation of $\tilde{e}$ can be precisely determined. Accordingly, the vector $\tilde{\mathbf{e}}$ for a group j was calculated as zero for group j and the optimal effort that the other group $-j$ would exert in the absence of group j. Participation constraints were calculated directly from (A.2). Note that with the two-group special case, it was not necessary to recalculate equilibrium participation decisions when one group's constraint fell below the threshold; with one group always choosing to enter the movement, the result is always a Nash equilibrium.

The three-dimensional graphs were the surface of the collected participation constraint matrix, with interpolated shading for different magnitudes. Both surfaces cross the threshold hyperplane at $PC = 0$, but only at sufficiently small values of $\hat{x}$. The code used for the simulations is available at the corresponding author's website.